

0400
04/10/2001

OIPE

#2

RAW SEQUENCE LISTING

DATE: 03/30/2001

PATENT APPLICATION: US/09/810,836

TIME: 15:09:46

Input Set : A:\MBI0032.txt

Output Set: N:\CRF3\03302001\I810836.raw

ENTERED

3 <110> APPLICANT: Broun, Pierre
 5 <120> TITLE OF INVENTION: Method for Modifying a Biosynthetic Pathway
 7 <130> FILE REFERENCE: MBI-0032
 C--> 9 <140> CURRENT APPLICATION NUMBER: US/09/810,836
 C--> 9 <141> CURRENT FILING DATE: 2001-03-16
 9 <160> NUMBER OF SEQ ID NOS: 8
 11 <170> SOFTWARE: PatentIn version 3.0
 13 <210> SEQ ID NO: 1
 14 <211> LENGTH: 1239
 15 <212> TYPE: DNA
 16 <213> ORGANISM: Arabidopsis thaliana
 18 <220> FEATURE:
 19 <221> NAME/KEY: CDS
 20 <222> LOCATION: (6)..(1091)
 21 <223> OTHER INFORMATION: G993
 24 <400> SEQUENCE: 1
 25 caaat atg gaa tac agc tgt gta gac gac agt agt aca acg tca gaa tct 50
 26 Met Glu Tyr Ser Cys Val Asp Asp Ser Ser Thr Thr Ser Glu Ser
 27 1 5 10 15
 29 ctg tcc atc tct act act cca aag ccg aca acg acg acg gag aag aaa 98
 30 Leu Ser Ile Ser Thr Thr Pro Lys Pro Thr Thr Thr Thr Glu Lys Lys
 31 20 25 30
 33 ctg tct tct ccg ccg gcg acg tgc atg cgt ctg tac aga atg gga agc 146
 34 Leu Ser Ser Pro Pro Ala Thr Ser Met Arg Leu Tyr Arg Met Gly Ser
 35 35 40 45
 37 ggc gga agc agc gtc gtt ttg gat tca gag aac ggc gtc gag acc gag 194
 38 Gly Gly Ser Ser Val Val Leu Asp Ser Glu Asn Gly Val Glu Thr Glu
 39 50 55 60
 41 tca cgt aag ctt cct tgc tgc aaa tat aaa ggc gtt gtg cct cag cct 242
 42 Ser Arg Lys Leu Pro Ser Ser Lys Tyr Lys Gly Val Val Pro Gln Pro
 43 65 70 75
 45 aac gga aga tgg gga gct cag att tac gag aag cat cag cga gtt tgg 290
 46 Asn Gly Arg Trp Gly Ala Gln Ile Tyr Glu Lys His Gln Arg Val Trp
 47 80 85 90 95
 49 ctg ggt act ttc aac gag gaa gaa gaa gct gcg tct tct tac gac atc 338
 50 Leu Gly Thr Phe Asn Glu Glu Glu Glu Ala Ala Ser Ser Tyr Asp Ile
 51 100 105 110
 53 gcc gtg agg aga ttc cgc gcc cgc gac gcc gtc act aac ttc aaa tct 386
 54 Ala Val Arg Arg Phe Arg Gly Arg Asp Ala Val Thr Asn Phe Lys Ser
 55 115 120 125
 57 caa gtt gat gga aac gac gcc gaa tgc gct ttt ctt gac gct cat tct 434
 58 Gln Val Asp Gly Asn Asp Ala Glu Ser Ala Phe Leu Asp Ala His Ser
 59 130 135 140
 61 aaa gct gag atc gtg gat atg ttg agg aaa cac act tac gcc gat gag 482
 62 Lys Ala Glu Ile Val Asp Met Leu Arg Lys His Thr Tyr Ala Asp Glu
 63 145 150 155
 65 ttt gag cag agt aga cgg aag ttt gtt aac ggc gac gga aaa cgc tct 530

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/810,836

DATE: 03/30/2001

TIME: 15:09:46

Input Set : A:\MBI0032.txt

Output Set: N:\CRF3\03302001\I810836.raw

```

66 Phe Glu Gln Ser Arg Arg Lys Phe Val Asn Gly Asp Gly Lys Arg Ser
67 160 165 170 175
69 ggg ttg gag acg gcg acg tac gga aac gac gct gtt ttg aga gcg cgt 578
70 Gly Leu Glu Thr Ala Thr Tyr Gly Asn Asp Ala Val Leu Arg Ala Arg
71 180 185 190
73 gag gtt ttg ttc gag aag act gtt acg ccg agc gac gtc ggg aag ctg 626
74 Glu Val Leu Phe Glu Lys Thr Val Thr Pro Ser Asp Val Gly Lys Leu
75 195 200 205
77 aac cgt tta gtg ata ccg aaa caa cac gcg gag aag cat ttt ccg tta 674
78 Asn Arg Leu Val Ile Pro Lys Gln His Ala Glu Lys His Phe Pro Leu
79 210 215 220
81 ccg gcg atg acg acg gcg atg ggg atg aat ccg tct ccg acg aaa ggc 722
82 Pro Ala Met Thr Thr Ala Met Gly Met Asn Pro Ser Pro Thr Lys Gly
83 225 230 235
85 gtt ttg att aac ttg gaa gat aga aca ggg aaa gtg tgg cgg ttc cgt 770
86 Val Leu Ile Asn Leu Glu Asp Arg Thr Gly Lys Val Trp Arg Phe Arg
87 240 245 250 255
89 tac agt tac tgg aac agc agt caa agt tac gtg ttg acc aag ggc tgg 818
90 Tyr Ser Tyr Trp Asn Ser Ser Gln Ser Tyr Val Leu Thr Lys Gly Trp
91 260 265 270
93 agc cgg ttc gtt aaa gag aag aat ctt cga gcc ggt gat gtg gtt tgt 866
94 Ser Arg Phe Val Lys Glu Lys Asn Leu Arg Ala Gly Asp Val Val Cys
95 275 280 285
97 ttc gag aga tca acc gga cca gac cgg caa ttg tat atc cac tgg aaa 914
98 Phe Glu Arg Ser Thr Gly Pro Asp Arg Gln Leu Tyr Ile His Trp Lys
99 290 295 300
101 gtc cgg tct agt ccg gtt cag act gtg gtt agg cta ttc gga gtc aac 962
102 Val Arg Ser Ser Pro Val Gln Thr Val Val Arg Leu Phe Gly Val Asn
103 305 310 315
105 att ttc aat gtg agt aac gag aaa cca aac gac gtc gca gta gag tgt 1010
106 Ile Phe Asn Val Ser Asn Glu Lys Pro Asn Asp Val Ala Val Glu Cys
107 320 325 330 335
109 gtt ggc aag aag aga tct cgg gaa gat gat ttg ttt tcg tta ggg tgt 1058
110 Val Gly Lys Lys Arg Ser Arg Glu Asp Asp Leu Phe Ser Leu Gly Cys
111 340 345 350
113 tcc aag aag cag gcg att atc aac atc ttg tga caaattcttt ttttttggtt 1111
114 Ser Lys Lys Gln Ala Ile Ile Asn Ile Leu
115 355 360
117 tttttcttca atttgtttct cctttttcaa tattttgtat tgaaatgaca agttgtaaat 1171
119 taggacaaga caagaaaaaa tgacaactag acaaaatagt ttttgtttaa aaaaaaaaaa 1231
121 aaaaaaaaaa 1239
124 <210> SEQ ID NO: 2
125 <211> LENGTH: 361
126 <212> TYPE: PRT
127 <213> ORGANISM: Arabidopsis thaliana
129 <400> SEQUENCE: 2
131 Met Glu Tyr Ser Cys Val Asp Asp Ser Ser Thr Thr Ser Glu Ser Leu
132 1 5 10 15
135 Ser Ile Ser Thr Thr Pro Lys Pro Thr Thr Thr Thr Glu Lys Lys Leu

```

RAW SEQUENCE LISTING

DATE: 03/30/2001

PATENT APPLICATION: US/09/810,836

TIME: 15:09:46

Input Set : A:\MBI0032.txt

Output Set: N:\CRF3\03302001\I810836.raw

```

136          20          25          30
139 Ser Ser Pro Pro Ala Thr Ser Met Arg Leu Tyr Arg Met Gly Ser Gly
140          35          40          45
143 Gly Ser Ser Val Val Leu Asp Ser Glu Asn Gly Val Glu Thr Glu Ser
144          50          55          60
147 Arg Lys Leu Pro Ser Ser Lys Tyr Lys Gly Val Val Pro Gln Pro Asn
148 65          70          75          80
151 Gly Arg Trp Gly Ala Gln Ile Tyr Glu Lys His Gln Arg Val Trp Leu
152          85          90          95
155 Gly Thr Phe Asn Glu Glu Glu Glu Ala Ala Ser Ser Tyr Asp Ile Ala
156          100         105         110
159 Val Arg Arg Phe Arg Gly Arg Asp Ala Val Thr Asn Phe Lys Ser Gln
160          115         120         125
163 Val Asp Gly Asn Asp Ala Glu Ser Ala Phe Leu Asp Ala His Ser Lys
164          130         135         140
167 Ala Glu Ile Val Asp Met Leu Arg Lys His Thr Tyr Ala Asp Glu Phe
168 145          150         155         160
171 Glu Gln Ser Arg Arg Lys Phe Val Asn Gly Asp Gly Lys Arg Ser Gly
172          165         170         175
175 Leu Glu Thr Ala Thr Tyr Gly Asn Asp Ala Val Leu Arg Ala Arg Glu
176          180         185         190
179 Val Leu Phe Glu Lys Thr Val Thr Pro Ser Asp Val Gly Lys Leu Asn
180          195         200         205
183 Arg Leu Val Ile Pro Lys Gln His Ala Glu Lys His Phe Pro Leu Pro
184          210         215         220
187 Ala Met Thr Thr Ala Met Gly Met Asn Pro Ser Pro Thr Lys Gly Val
188 225          230         235         240
191 Leu Ile Asn Leu Glu Asp Arg Thr Gly Lys Val Trp Arg Phe Arg Tyr
192          245         250         255
195 Ser Tyr Trp Asn Ser Ser Gln Ser Tyr Val Leu Thr Lys Gly Trp Ser
196          260         265         270
199 Arg Phe Val Lys Glu Lys Asn Leu Arg Ala Gly Asp Val Val Cys Phe
200          275         280         285
203 Glu Arg Ser Thr Gly Pro Asp Arg Gln Leu Tyr Ile His Trp Lys Val
204          290         295         300
207 Arg Ser Ser Pro Val Gln Thr Val Val Arg Leu Phe Gly Val Asn Ile
208 305          310         315         320
211 Phe Asn Val Ser Asn Glu Lys Pro Asn Asp Val Ala Val Glu Cys Val
212          325         330         335
215 Gly Lys Lys Arg Ser Arg Glu Asp Asp Leu Phe Ser Leu Gly Cys Ser
216          340         345         350
219 Lys Lys Gln Ala Ile Ile Asn Ile Leu
220          355         360
223 <210> SEQ ID NO: 3
224 <211> LENGTH: 1226
225 <212> TYPE: DNA
226 <213> ORGANISM: Arabidopsis thaliana
228 <220> FEATURE:
229 <221> NAME/KEY: CDS

```

RAW SEQUENCE LISTING

DATE: 03/30/2001

PATENT APPLICATION: US/09/810,836

TIME: 15:09:46

Input Set : A:\MBI0032.txt

Output Set: N:\CRF3\03302001\I810836.raw

```

230 <222> LOCATION: (111)..(989)
231 <223> OTHER INFORMATION: G1845
234 <400> SEQUENCE: 3
235 aagacataat tttctctgtt ttcctagctc tctcctctca aattcttcca ttgctctctg      60
237 ttttggcaaa tcgtgaactg ccacgtcttt aaggcatcag tgaagcaaaag atg gac      116
238                                     Met Asp
239                                     1
241 ttt gac gag gag cta aat ctt tgt att acg aaa ggt aaa aat gtt gat      164
242 Phe Asp Glu Glu Leu Asn Leu Cys Ile Thr Lys Gly Lys Asn Val Asp
243      5      10      15
245 cat tct ttt gga gga gaa gct tct tcc acg tcc cca aga tct atg aag      212
246 His Ser Phe Gly Gly Glu Ala Ser Ser Thr Ser Pro Arg Ser Met Lys
247      20      25      30
249 aaa atg aag agt cct agt cgt cct aaa ccc tat ttc caa tcc tct tct      260
250 Lys Met Lys Ser Pro Ser Arg Pro Lys Pro Tyr Phe Gln Ser Ser Ser
251 35      40      45      50
253 tct cct tat tcg tta gag gct ttc cct ttt tct ctc gat cca aca ctt      308
254 Ser Pro Tyr Ser Leu Glu Ala Phe Pro Phe Ser Leu Asp Pro Thr Leu
255      55      60      65
257 cag aat cag caa caa caa ctc gga tca tac gtt ccg gta ctt gag caa      356
258 Gln Asn Gln Gln Gln Gln Leu Gly Ser Tyr Val Pro Val Leu Glu Gln
259      70      75      80
261 cga caa gac ccg aca atg caa ggc cag aag caa atg atc tcc ttt agt      404
262 Arg Gln Asp Pro Thr Met Gln Gly Gln Lys Gln Met Ile Ser Phe Ser
263      85      90      95
265 cct caa caa caa caa cag cag cag cag tat atg gcc cag tac tgg agt      452
266 Pro Gln Gln Gln Gln Gln Gln Gln Gln Tyr Met Ala Gln Tyr Trp Ser
267 100      105      110
269 gac aca ttg aat ctg agt cca aga gga aga atg atg atg atg atg agc      500
270 Asp Thr Leu Asn Leu Ser Pro Arg Gly Arg Met Met Met Met Ser
271 115      120      125      130
273 caa gaa gct gtt caa cct tac atc gca acg aag ctg tac aga gga gtg      548
274 Gln Glu Ala Val Gln Pro Tyr Ile Ala Thr Lys Leu Tyr Arg Gly Val
275      135      140      145
277 aga caa cgt caa tgg gga aaa tgg gtc gca gag atc cgt aag cca cga      596
278 Arg Gln Arg Gln Trp Gly Lys Trp Val Ala Glu Ile Arg Lys Pro Arg
279      150      155      160
281 agc agg gca cgt ctt tgg ctt ggt acc ttt gat aca gct gaa gaa gct      644
282 Ser Arg Ala Arg Leu Trp Leu Gly Thr Phe Asp Thr Ala Glu Glu Ala
283      165      170      175
285 gcc atg gcc tac gac cgc caa gcc ttc aaa tta cga ggc cac agc gca      692
286 Ala Met Ala Tyr Asp Arg Gln Ala Phe Lys Leu Arg Gly His Ser Ala
287      180      185      190
289 aca ctg aat ttc ccg gag cat ttt gtg aat aag gaa agc gag ctg cat      740
290 Thr Leu Asn Phe Pro Glu His Phe Val Asn Lys Glu Ser Glu Leu His
291 195      200      205      210
293 gat tca aac tcg tcg gat cag aaa gaa cct gaa acg cca cag cca agc      788
294 Asp Ser Asn Ser Ser Asp Gln Lys Glu Pro Glu Thr Pro Gln Pro Ser
295      215      220      225

```

RAW SEQUENCE LISTING

DATE: 03/30/2001

PATENT APPLICATION: US/09/810,836

TIME: 15:09:46

Input Set : A:\MBI0032.txt

Output Set: N:\CRF3\03302001\I810836.raw

```

297 gag gtt aac ttg gag agc aag gaa cta ccg gtg att gat gtt ggg aga      836
298 Glu Val Asn Leu Glu Ser Lys Glu Leu Pro Val Ile Asp Val Gly Arg
299          230          235          240
301 gag gaa ggt atg gct gag gca tgg tac aat gcc att aca tcg gga tgg      884
302 Glu Glu Gly Met Ala Glu Ala Trp Tyr Asn Ala Ile Thr Ser Gly Trp
303          245          250          255
305 ggt cct gaa agt cct ctt tgg gat gat ttg gat agt tct cat cag ttt      932
306 Gly Pro Glu Ser Pro Leu Trp Asp Asp Leu Asp Ser Ser His Gln Phe
307          260          265          270
309 tca tca gaa agc tca tct tct tct cct ctc tct tgt cct atg agg cct      980
310 Ser Ser Glu Ser Ser Ser Ser Ser Pro Leu Ser Cys Pro Met Arg Pro
311 275          280          285          290
313 ttc ttt tga aaaagtttat aaaccacat tgtgtttag gttatagttt      1029
314 Phe Phe
317 agggttatgc tcattggcat ttggatggag gcaatttttg tgatctccca ttccaccaca      1089
319 tatcagtcac tatatgtgtc taccttttct ctgtatttct atcattatca ttgtttttat      1149
321 tatgtgtctg tatgtgtttc cctattgcta catacataga tgcctcttt gttcaaaaaa      1209
323 aaaaaaaaaa aaaaaaa      1226
326 <210> SEQ ID NO: 4
327 <211> LENGTH: 292
328 <212> TYPE: PRT
329 <213> ORGANISM: Arabidopsis thaliana
331 <400> SEQUENCE: 4
333 Met Asp Phe Asp Glu Glu Leu Asn Leu Cys Ile Thr Lys Gly Lys Asn
334 1          5          10          15
337 Val Asp His Ser Phe Gly Gly Glu Ala Ser Ser Thr Ser Pro Arg Ser
338          20          25          30
341 Met Lys Lys Met Lys Ser Pro Ser Arg Pro Lys Pro Tyr Phe Gln Ser
342          35          40          45
345 Ser Ser Ser Pro Tyr Ser Leu Glu Ala Phe Pro Phe Ser Leu Asp Pro
346          50          55          60
349 Thr Leu Gln Asn Gln Gln Gln Gln Leu Gly Ser Tyr Val Pro Val Leu
350 65          70          75          80
353 Glu Gln Arg Gln Asp Pro Thr Met Gln Gly Gln Lys Gln Met Ile Ser
354          85          90          95
357 Phe Ser Pro Gln Gln Gln Gln Gln Gln Gln Tyr Met Ala Gln Tyr
358          100          105          110
361 Trp Ser Asp Thr Leu Asn Leu Ser Pro Arg Gly Arg Met Met Met Met
362          115          120          125
365 Met Ser Gln Glu Ala Val Gln Pro Tyr Ile Ala Thr Lys Leu Tyr Arg
366          130          135          140
369 Gly Val Arg Gln Arg Gln Trp Gly Lys Trp Val Ala Glu Ile Arg Lys
370 145          150          155          160
373 Pro Arg Ser Arg Ala Arg Leu Trp Leu Gly Thr Phe Asp Thr Ala Glu
374          165          170          175
377 Glu Ala Ala Met Ala Tyr Asp Arg Gln Ala Phe Lys Leu Arg Gly His
378          180          185          190
381 Ser Ala Thr Leu Asn Phe Pro Glu His Phe Val Asn Lys Glu Ser Glu
382          195          200          205

```

VERIFICATION SUMMARY

DATE: 03/30/2001

PATENT APPLICATION: US/09/810,836

TIME: 15:09:47

Input Set : A:\MBI0032.txt

Output Set: N:\CRF3\03302001\I810836.raw

L:9 M:270 C: Current Application Number differs, Replaced Current Application No
L:9 M:271 C: Current Filing Date differs, Replaced Current Filing Date